Attorney's Docket No.:10559/195001/P8367/Intel Corporation

In the claims:

1. (Currently amended)   A method comprising:

segmenting video data to create a video clip based on

timing data that indicates a specified timing within a gesture

will occur; and

determining information related to a a most likely gesture

occurring in the video clip only at the specified timing.

2. (Currently amended) The method of claim 1, wherein

determining includes determining a probability that each of a

plurality of predefined gestures which are performed in the

video clip contains the predefined gesture.

3. (Original)   The method of claim 2, wherein determining

the probability that the video clip contains each of the

predefined gesture includes evaluations of Hidden Markov Models.

4.   (Original) The method of claim 1, wherein the timing

data includes beat data corresponding to a beat of audio data.

5. (Original)   The method of claim 4, further comprising:

receiving the audio data; and

extracting the beat data from the audio data.

6

6. (Original) The method of claim 4, wherein the video clip includes a portion of the video data corresponding to a predefined time window surrounding the occurrence of at least one beat.

7. (Original) The method of claim 1, further comprising displaying a target gesture to be performed by the subject of the video data.

8. (Original) The method of claim 1, wherein each video clip contains video frames.

9. (Original) The method of claim 1, further comprising identifying moving regions in each video frame in the video clip.

10. (Original) The method of claim 9, further comprising generating a feature vector for each video frame of the video clip.

11. (Original) The method of claim 1, further comprising generating a score based on whether the video clip contains the target gesture.

7

Attorney's Docket No.:10559/195001/P8367/Intel Corporation

12. (Original) The method of claim 11, further comprising displaying the score.

13. (Original) The method of claim 1, wherein determining if the video clip contains the predefined gesture includes generating a gesture probability vector having a plurality of elements, each element being associated with one of a plurality of predefined gestures and representing a probability that the video clip contains each of the associated predefined gestures.

14. (Currently amended) A system comprising:

a temporal segmentor connected to receive video data and to create a video clip from the video data based on timing data that indicates a specified timing within which a gesture will occur; and

a recognition engine, in communication with the temporal segmentor, to determine if the video clip contains a predefined gesture, only at the specified timing.

15. (Original) The system of claim 14, wherein the recognition engine includes a plurality of Hidden Markov Models.

16. (Original) The system of claim 14, further comprising:

8

a timing data source, in communication with the temporal

segmentor, to provide the timing data to the temporal segmentor;

and

a video source, in communication with the temporal

segmentor, to provide the video data to the temporal segmentor.

17. (Original) The system of claim 14, further comprising a

move subsystem, in communication with the timing data source, to

provide a target gesture to be performed by the subject of the

video data.

18. (Original) The system of claim 17, wherein the target

gesture is a dance move that is to be performed by the subject

of the video data.

19. (Original) The system of claim 17, further comprising a

scoring subsystem, in communication with the recognition engine

and the move subsystem, to determine if the video clip contains

the target gesture.

20.(Original) The system of claim 19, further comprising a

display subsystem, in communication with the scoring subsystem,

to display a score that is a function of whether the video clip

contains the target gesture.

21. (Original) The system of claim 20, wherein the display subsystem is in communication with the move subsystem and is configured to display a gesture request based on the target gesture.

22. (Original) The system of claim 14, wherein the recognition engine is configured to recognize predefined gestures and to produce a gesture probability vector having elements, each element being associated with one of the predefined gestures and representing the probability that the video clip contains the associated predefined gesture.

23. (Original) The system of claim 14, wherein the timing data source includes:

an audio source that provides an audio data; and

a beat extractor, in communication with the audio source, that extracts beat data from the audio data.

24. (Original) The system of claim 23, wherein the video clip corresponds to a beat in the beat data.

10

25. (Original) The system of claim 24, wherein the video clip includes a portion of the video data corresponding to a predefined time window surrounding the occurrence of the beat.

26. (Currently amended) A computer program product, tangibly stored on a computer-readable medium, for recognizing gestures contained in video data, comprising instructions operable to cause a programmable processor to:

segment the video data to create a video clip based on timing data <u>that indicates a specified timing within which a gesture will occur</u>; and

determine if the video clip contains a predefined gesture <u>within the specified timing</u>.

27. (Original) The product of claim 26, further comprising instructions operable to cause the programmable processor to:

extract beat data from an audio signal; and

segment the video data to create the video clip using the beat data.

28. (Currently amended) An audio-visual processing system including:

a video source to provide video data;

an audio source to provide audio data;

11

a speaker to play at least a portion of the audio data; and

a computer program product, tangibly stored on a computer-readable medium, for recognizing gestures contained in video data, comprising instructions operable to cause a programmable processor, in communication with the video source and the audio source, to:

extract beat data from the audio data;

segment the video data to create a video clip based on said beat data; and

determine if the video clip contains a predefined gesture within only a specified timing related to said beat data.

29. (Original) The video processing system of claim 28, wherein the computer program product further includes instructions operable to cause the programmable processor to:

perform a Hidden Markov Model process to determine if the video clip contains the predefined gesture.

30. (Original) The video processing system of claim 28, further comprising a display to display information based on whether the video clip contains the predefined gesture.

12